ON THE USE OF AN INDEX OF BIAS IN THE ESTIMATION OF POWER SPECTRA

HIROTUGU AKAIKE

(Received Oct. 26, 1967)

1. Introduction

In the practical situation of estimation of a power spectrum we shall proceed as follows:

1. First, we set the resolvability required of the estimate, which determines the bandwidth (in some sense or other) of the spectrum window to be adopted for estimation, and then

2. We settle the requirement of the statistical accuracy or the sampling variability of the estimate, which is determined by the length of data used in the computation.

In case these two requirements of the estimate are given it is a simple matter to provide an estimate satisfying these two conditions, if only the desired length of observation is available. But, in most of the practical situations of estimation of power spectra, we usually can not have much confidence in the primary selection of resolvability of the estimate (or the bandwidth of the spectrum window) and it is a common practice to evaluate the adequacy of the primary selection of resolvability after the result of computation of the spectrum is obtained. Some people consider that we can evade the difficulty by setting the required resolvability high enough. Unfortunately it is usual that such a selection of resolvability requires formidable length of original data to attain the desired low level of sampling variability. This means that our smoothing operator (spectrum window) in the frequency domain usually has to have a fairly narrow pass band (or a wide bandwidth of spectrum window) to suppress the noise due to sampling fluctuation and thus we inevitably have to pay our attention to the linear distortion, or the bias, introduced by applying the smoothing operator. The most natural way of evaluation of this distortion would be to change the window in an intended manner so as to make the difference of estimates sensitive to a prescribed type of bias and check the difference against its expected sampling variability. This is the procedure we are going to discuss in this paper, and it is in close relation with the procedure proposed by H. E. Daniell [4].

There is another point to which we must pay our attention. This is the practicability of the procedure, or the amount of necessary computations for the procedure. The procedure is kept most simple and economical if we adopt spectrum windows corresponding to lag windows of trigonometric sum type [1, 3]. We shall see that the lag window with different bias characteristics, obtained by the present author [1], can quite conveniently be used for the present purpose. Applications to an artificial set of data and two sets of real data are illustrated, which suggest the practical applicability of the procedure. A result of a Monte Carlo experiment is also included to verify the validity of mathematical derivations in this article.

2. Test statistics or an index of bias

Throughout the present paper, we shall assume that a record $\{x(t): -T<t<T\}$ of stationary time series is given and consider that type of estimate of power spectrum which is given by a smoothed periodogram. Thus our estimate would be of the form

$$\hat{p}_w(f_j)=\sum_k W_k I(f_j-f_k),$$

where

$$I(f_j)=\left|\frac{1}{\sqrt{2T}}\int_{-T}^{T}\exp\left(-i2\pi\frac{j}{2T}t\right)x(t)dt\right|^2,$$

$$f_k=\frac{k}{2T} \qquad (k=0, \pm 1, \pm 2, \cdots)$$

and $\{W_k\}$ is a set of real numbers satisfying $\sum_k W_k=1$ and defining the smoothing operator or the spectrum window. When the original data $x(t)$ is discrete in time, the integration in the above formula should be interpreted as a corresponding summation.

We define our index of bias $b(f_j)$ by

$$b(f_j)=\frac{\hat{p}_{\Delta w}(f_j)}{\hat{p}_w(f_j)},$$

where $\hat{p}_{\Delta w}(f_j)$ is obtained by replacing $\{W_k\}$ in the definition of $\hat{p}_w(f_j)$ by $\{\Delta W_k\}$ satisfying the condition $\sum_k \Delta W_k=0$.

We select $\{W_k\}$ as an ordinary spectrum window which responds to rather slow (low frequency) change of the spectrum in frequency and $\{\Delta W_k\}$ as a window which responds mainly to the rapid (high frequency) change of the spectrum. Thus if there is a significant energy of power spectrum at higher frequencies (in frequency) filtered out by $\{W_k\}$, $b(f_j)$ will show a significantly positive or negative value and otherwise it will remain in the neighborhood of zero. It is expected that $\{\Delta W_k\}$ can be such that $b(f_j)$ will be significantly positive where $\hat{p}_w(f_j)$ is significantly lower than the true $p(f_j)$ and vice versa. We should have only to check the value of $b(f_j)$ against its expected sampling variability to evaluate the significance of the value.

In this article we assume that $I(f_j)$'s appearing in the summation are mutually independent and following negative exponential distribution with $EI(f_j)=p(f_j)$, where $p(f)$ is the power spectral density function of the process $x(t)$ and is assumed to be locally sufficiently smooth to make the present assumption reasonable (see section 4 of [2])*.

We have

$$E\hat{p}_w(f_j)=\sum_k W_k p(f_j-f_k)$$

and

$$D^2\hat{p}_w(f_j)=\sum_k W_k^2 p^2(f_j-f_k)\,.$$

Thus if we can assume in the summation that $p(f_j-f_k)$ is nearly equal to $p(f_j)$ we get the (approximate) relations

$$E\hat{p}_w(f_j)=p(f_j)$$

$$D^2\hat{p}_w(f_j)=(\sum_k W_k^2)p^2(f_j)\,.$$

Now we assume that $\{W_k\}$ is the set of Fourier coefficients of a square integrable time-domain function $D(t)$, i.e., W_k is given by the formula

$$W_k=\frac{1}{2T}\int_{-T}^{T}\exp\left(-i2\pi\frac{k}{2T}\right)D(t)dt\,.$$

Then we have

$$\sum_k W_k^2=\frac{1}{2T}\int_{-T}^{T}D^2(t)dt\,.$$

Thus if we fix a $D(t)$ which vanishes outside a finite interval of t and make T large so that $D(t)=0$ for $|t|>T$ we get the relation

$$\sum_k W_k^2=\frac{1}{2T}\int_{-\infty}^{\infty}D^2(t)dt\,.$$

This shows that the variance of $\hat{p}_w(f_j)$ is inversely proportional to $2T$ (the observation length). In practical situations of spectrum analysis

* In practical situation $I(-f_i)=I(f_i)$ holds. We are assuming in the following discussion that the contribution from the negative frequencies is neglegibly small. Also we are assuming that the mean of the process under observation is vanishing.

we usually keep $D^2\hat{p}_w(f_j)/(E\hat{p}_w(f_j))^2$ fairly smaller than 1, say less than 1/10, and thus it would be useful to evaluate the statistical properties of a statistics

$$\beta(f_j)=\frac{\hat{p}_{\Delta w}(f_j)}{p(f_j)}\left(1-\frac{\hat{p}_w(f_j)-p(f_j)}{p(f_j)}\right)$$

which is an approximation to $b(f_j)$ when $|(\hat{p}_w(f_j)-p(f_j))/p(f_j)|$ is sufficiently smaller than 1.

Under the present assumption on the distribution of $I(f_\nu)$'s we can easily show that the following relations hold:

$$E\beta(f_j)=-\sum_k (\Delta W_k)W_k\,,$$

$$D^2\beta(f_j)=\sum_k (\Delta W_k)^2-4\sum_k (\Delta W_k)^2 W_k+6\sum_k (\Delta W_k)^2 W_k^2 +(\sum_k (\Delta W_k)^2)(\sum_k W_k^2)+(\sum_k (\Delta W_k)W_k)^2\,.$$

We can use these quantities to evaluate approximately the sampling variability of our index of bias $b(f_j)$. When $\{W_k\}$ and $\{\Delta W_k\}$ are given we can directly calculate $E\beta(f_j)$ and $D^2\beta(f_j)$ by using the above formulae. But as will be seen in the next section, the computation of necessary quantities is quite simple when we adopt as $D(t)$ a lag window of trigonometric sum type.

3. Evaluation of $E\beta$ and $D^2\beta$ for a window of trigonometric sum type

Here we consider the case where $D(t)$ corresponding to $\{W_k\}$ is given as follows:

$$D(t)=\begin{cases}\sum_{n=-K}^{K} a_n \exp\left(-i2\pi\frac{n}{2L}t\right) & \text{for } |t|\leqq L\,,\\ 0 & \text{otherwise}\,,\end{cases}$$

where it is assumed that $L<T$ and that a_n is real, $a_{-n}=a_n$ and $\sum_n a_n=1$.

We shall also assume that $\{\Delta W_k\}$ corresponds to a difference of two time domain functions of this type, i.e., $\{\Delta W_k\}$ is the set of Fourier coefficients of a function $\Delta D(t)$ which is given by

$$\Delta D(t)=\begin{cases}\sum_{n=-K}^{K} \Delta a_n \exp\left(-i2\pi\frac{n}{2L}t\right) & \text{for } |t|\leqq L\,,\\ 0 & \text{otherwise}\,,\end{cases}$$

where Δa_n is real, $\Delta a_{-n}=\Delta a_n$ and $\sum_n \Delta a_n=0$.

In this case we have the following identities:

$$\sum_k (\Delta W_k) W_k = \frac{2L}{2T} \sum_{n=-K}^{K} (\Delta a_n) a_n ,$$

$$\sum_k W_k^2 = \frac{2L}{2T} \sum_{n=-K}^{K} a_n^2 ,$$

$$\sum_k (\Delta W_k)^2 = \frac{2L}{2T} \sum_{n=-K}^{K} (\Delta a_n)^2 ,$$

$$\sum_k (\Delta W_k)^2 W_k = \left(\frac{2L}{2T}\right)^2 \left[\frac{3}{4} \sum_{n=-K}^{K} (\Delta a_n)^2 a_n + \frac{2}{(2\pi)^2} \sum_{n=-K}^{K} ((\Delta a_n)^2 z_n - \gamma_n y_n)\right] ,$$

$$\sum_k (\Delta W_k)^2 W_k^2 = \left(\frac{2L}{2T}\right)^3 \frac{2}{(2\pi)^2} \left[\frac{4}{3} \sum_{n=-K}^{K} (\Delta a_n)^2 a_n^2 + 2 \sum_{n=-K}^{K} ((\Delta a_n)^2 \zeta_n - \gamma_n \eta_n + \delta_n \xi_n - \gamma_n \delta_n)\right] ,$$

where

$$y_n = \sum_{m \neq n} a_m \frac{1-(-1)^{n-m}}{(n-m)}$$

$$z_n = \sum_{m \neq n} a_m \frac{1-(-1)^{n-m}}{(n-m)^2}$$

$$\gamma_n = \left[\sum_{m \neq n} \Delta a_m \frac{(-1)^{n-m}}{(n-m)}\right] \Delta a_n$$

$$\delta_n = \left[\sum_{m \neq n} a_m \frac{(-1)^{n-m}}{(n-m)}\right] a_n$$

$$\xi_n = \sum_{m \neq n} (\Delta a_m)^2 \frac{1}{(n-m)}$$

$$\eta_n = \sum_{m \neq n} a_m^2 \frac{1}{(n-m)}$$

$$\zeta_n = \sum_{m \neq n} a_m^2 \frac{1}{(n-m)^2}$$

and the summation $\sum_{m \neq n}$ is extended all over the m's which are different from n.

Thus if we are given a properly chosen set of two trigonometric sum type lag windows $D_1(t)$ and $D_2(t)$ and put $D(t) = D_1(t)$ and $\Delta D(t) = D_2(t) - D_1(t)$ we can easily evaluate $E\beta(f_j)$ and $D^2\beta(f_j)$ by using the present result. We can see how the lag window of trigonometric sum type is suited to get an overall view of the sampling variabilities of related quantities. Also we should pay our attention to the computational ease obtained by the adoption of this type of windows [1, 3].

Now, we will give a concrete numerical example. In the paper [1] cited above, various kinds of windows of trigonometric sum type were constructed. Here we put $K=2$ and adopt as $D_1(t)$ $(=D(t))$ the one given by the set of coefficients $(a_0=0.5132,\ a_1=a_{-1}=0.2434,\ a_2=a_{-2}=0)$, which is bias free for local variation of $p(f)$ of odd orders, and as $D_2(t)$ $(=D(t)+\Delta D(t))$ the one given by

$$(a_0=0.6398,\ a_1=a_{-1}=0.2401,\ a_2=a_{-2}=-0.0600)$$

which is expected to be bias free for locally quadratic variation of $p(f)$ besides that of odd orders. These lag windows were denoted in the paper as $W_1(0, \alpha\beta)$ and $W_2(2, \alpha\beta)$, respectively.

We have, in this case,

$$\sum_k W_k^2 = 0.763723\left(\frac{L}{2T}\right)$$

$$\sum_k (\Delta W_k) W_k = 0.126849\left(\frac{L}{2T}\right)$$

$$\sum_k (\Delta W_k)^2 = 0.046499\left(\frac{L}{2T}\right)$$

$$\sum_k (\Delta W_k)^2 W_k = 0.029849\left(\frac{L}{2T}\right)^2$$

$$\sum_k (\Delta W_k)^2 W_k^2 = 0.027584\left(\frac{L}{2T}\right)^3$$

$$\sum_k (\Delta W_k)^2 (\sum_k W_k^2) = 0.035512\left(\frac{L}{2T}\right)^2$$

$$(\sum_k (\Delta W_k) W_k)^2 = 0.016060\left(\frac{L}{2T}\right)^2 .$$

Thus we get

$$E\beta(f_j) = -0.126849\left(\frac{L}{2T}\right)$$

$$D^2\beta(f_j) = 0.046499\left(\frac{L}{2T}\right)\left(1-1.4586\left(\frac{L}{2T}\right)+3.5593\left(\frac{L}{2T}\right)^2\right).$$

As a rough approximation we have, assuming $\frac{L}{2T}$ to be small,

$$D\beta(f_j) = 0.216\sqrt{\frac{L}{2T}}\left(1-0.73\left(\frac{L}{2T}\right)+1.51\left(\frac{L}{2T}\right)^2\right).$$

Thus if $\frac{L}{2T}=\frac{1}{10}$ we have $E\beta(f_j)=-0.013$ and $D\beta(f_j)=0.064$. When $\frac{L}{2T}<\frac{1}{10}$ holds we can practically consider $E\beta(f_j)$ to be almost equal to zero and $D\beta(f_j)$ to be nearly equal to $0.216\sqrt{\frac{L}{2T}}$. This is equivalent to replacing $\beta(f_j)$ by a much simpler $\frac{\hat{p}_{\Delta w}(f_j)}{p(f_j)}$ to neglect the contribution of sampling fluctuation of the denominator.

It should be noted that, in practical situation, our present result has to be modified slightly at and around the frequency $f_j=0$ and $f_j=$ folding frequency ($=\frac{1}{2\Delta t}$: Δt=time interval between observations of original data or a sample covariance function). At these frequencies, there are usually various kinds of difficulties we shall not discuss the point in any detail here.

4. A result of Monte Carlo experiments

It would be interesting to verify the results of section 3 by Monte Carlo experiments. For this purpose we have taken 10 records of outputs of a random number generator each of length of 1000. The generator was developed by Mr. M. Isida of the Institute of Statistical Mathematics and is supposed to produce a 6 bits uniformly distributed random number at each call from the main computer. $\hat{p}_{w_1}(f_j)$ and $\hat{p}_{w_2}(f_j)$ were computed by first taking the (non-circular) mean lagged product, truncating it at lag 100, taking Fourier transform and then averaging by the weights given in section 3.

Computations were limited to the frequencies $f_j=\frac{\nu}{200}$ ($\nu=0, 1, 2, \cdots, 100$), where the time intervals between observations are assumed to be equal to unity. To avoid the possible boundary effects on the distribution of $\hat{p}_{w_1}(f_j)$ and $\hat{p}_{w_2}(f_j)$ we have further limited our observation to the frequency range $\{f_j=\frac{\nu}{200};\ \nu=6, 7, \cdots, 95\}$. Thus for each record we have 90 $b(f_j)$'s and we have 900 $b(f_j)$'s in all. Obviously there is a dependence between $b(f_j)$s within a record.

We observed the distributions of $b\left(\frac{\nu}{200}\right)$'s in each record. Apparently there was no significant variation of the distribution between the records, and we gathered all the $b\left(\frac{\nu}{200}\right)$'s ($\nu=6, 7, \cdots, 95$) of 10 records to get a distribution of $b(f_j)$ of size 900. This distribution is shown in Table 1 with its (sample) mean and standard deviation of ungrouped data. These two statistics are compared with the corresponding theoretical values of $E\beta(f_j)$ and $D\beta(f_j)$ obtained by the formulae of section 3. We can see a fairly good agreement between the corresponding quantities. It thus seems that $E\beta(f_j)$ and $D\beta(f_j)$ are useful approximations to $Eb(f_j)$ and $Db(f_j)$, respectively, for this value $\frac{1}{10}$ of $\frac{L}{2T}$.

Table 1

$b(f)$	frequency
−0.620	1
−0.340	1
−0.260	1
−0.220	10
−0.180	16
−0.140	35
−0.100	83
−0.060	137
−0.020	180
0.020	218
0.060	159
0.100	53
0.140	5
0.180	1

Sample mean $=-0.0128$
Sample standard deviation $=0.0724$
$E\beta(f_j)=-0.0127$
$D\beta(f_j)=0.0643$

Table 1 also shows a skewness of the distribution of $b(f_j)$ with a longer negative tail. This suggests that when we adopt the windows of section 3 we should expect that highly negative values of $b(f_j)$ could appear as a result of mere sampling fluctuation. Thus we should be more tolerate in deciding the significance of a negative value by scaling with their standard deviation than in the case of a positive value. It was observed that significantly negative or positive values of $b(f_j)$ appeared at those f_j where $\hat{p}_{w_1}(f_j)$ showed low or high values respectively. This shows the positive correlation between $\hat{p}_{\Delta w}(f_j)$ and $\hat{p}_w(f_j)$ and explains why $E\beta(f_j)$ is negative in this case.

The results of observations of this section should be taken into account in understanding the meaning of the numerical examples of the following section 5.

5. Numerical examples

We can most clearly see by numerical examples how our $D_1(t)$ and $D_2(t)$ of section 3 work in practical situations. As the first example, we computed $b(f_j)$'s by applying $D_1(t)$ and $D_2(t)$ to a theoretical covariance function which was free from sampling fluctuations. The results of computation are given in tables 2 and 3. The difference between $D_1(t)$ and $D_2(t)$ can be clearly seen and we can get some feeling about the behavior of $b(f_j)$ at the frequencies where $p(f)$ shows a rapid variation. Especially we can see that $b(f_j)$ shows significantly positive values at the peak of $p(f)$ and significantly negative values at its knees. This tendency is also clearly observed in the applications to real data. In tables 4, 5, 6 and 7 are given some results of applications to the analysis of random vibrations of an automobile. In these examples,

Table 2

f_j	$\hat{p}_{w_1}(f_j)$	$b(f_j)$	$C(f_j)$
$\frac{9}{100}$	0.5289	−0.033	−0.031
$\frac{10}{100}$	0.7141	−0.057	−0.043
$\frac{11}{100}$	1.0639	−0.088	−0.074
$\frac{12}{100}$	1.8019	−0.201*	−0.115
$\frac{13}{100}$	4.0124	−0.195*	−0.226
$\frac{14}{100}$	9.4545	+0.046	−0.142
$\frac{15}{100}$	13.1725	+0.141*	+1.444
$\frac{16}{100}$	9.1050	+0.041	−0.162
$\frac{17}{100}$	3.5728	−0.230*	−0.260
$\frac{18}{100}$	1.4309	−0.253*	−0.139
$\frac{19}{100}$	0.7639	−0.116*	−0.096
$\frac{20}{100}$	0.4623	−0.081	−0.059
$\frac{21}{100}$	0.3102	−0.048	−0.045

Note:

For the purpose of illustration, we assume $\frac{L}{2T}=\frac{1}{20}$, and we have

$$E\beta(f_j)=-0.006$$
$$D\beta(f_j)=0.047\,.$$

$$C(f_j)=\frac{p(f_j)-\hat{p}_{w_1}(f_j)}{\hat{p}_{w_1}(f_j)}$$

where

$p(f_j)$ is the theoretical value.

In tables 2 and 3 the unit of time is taken to be equal to 1.

In tables 2 through 7 "*" shows that $|b(f_j)-E\beta(f_j)|>2D\beta(f_j)$ holds.

Table 3

f_j	$\hat{p}_{w_1}(f_j)$	$b(f_j)$	$C(f_j)$
$\frac{24}{200}$	1.6449	−0.034	−0.030
$\frac{25}{200}$	2.2535	−0.046	−0.041
$\frac{26}{200}$	3.2910	−0.064	−0.056
$\frac{27}{200}$	5.1971	−0.079	−0.075
$\frac{28}{200}$	8.7874	−0.041	−0.077
$\frac{29}{200}$	13.8835	+0.042	+0.021
$\frac{30}{200}$	16.5869	+0.082	+0.147
$\frac{31}{200}$	13.6349	+0.042	+0.021
$\frac{32}{200}$	8.3551	−0.047	−0.086
$\frac{33}{200}$	4.7097	−0.091	−0.086
$\frac{34}{200}$	2.8320	−0.076	−0.066
$\frac{35}{200}$	1.8426	−0.056	−0.049
$\frac{36}{200}$	1.2793	−0.041	−0.037

We assume $\frac{L}{2T}=\frac{1}{10}$, and we have

$$E\beta(f_j)=-0.013$$
$$D\beta(f_j)=0.064\,.$$

Table 4

$f_j \Delta t$	$\hat{p}_{w_1}(f_j)$	$b(f_j)$
$\frac{3}{50}$	9.4240	+0.025
$\frac{4}{50}$	4.9637	−0.538*
$\frac{5}{50}$	11.5907	−0.334*
$\frac{6}{50}$	34.1257	+0.068
$\frac{7}{50}$	47.3913	+0.160*
$\frac{8}{50}$	28.7036	+0.015
$\frac{9}{50}$	8.6252	−0.434*
$\frac{10}{50}$	3.9897	−0.236*
$\frac{11}{50}$	3.1080	−0.267*
$\frac{12}{50}$	3.0531	−0.001
$\frac{13}{50}$	2.9555	−0.014

$2T=781\Delta t$ $E\beta(f_j)=-0.004$

$L=25\Delta t$ $D\beta(f_j)=0.038$

Table 5

$f_j \Delta t$	$\hat{p}_{w_1}(f_j)$	$b(f_j)$
$\frac{6}{100}$	11.8593	+0.091
$\frac{7}{100}$	4.2183	−0.276*
$\frac{8}{100}$	1.5690	−0.763*
$\frac{9}{100}$	3.7571	−0.206*
$\frac{10}{100}$	9.1256	−0.052
$\frac{11}{100}$	16.5294	−0.114*
$\frac{12}{100}$	31.4081	−0.055
$\frac{13}{100}$	53.8570	+0.070
$\frac{14}{100}$	60.3067	+0.078
$\frac{15}{100}$	47.8660	+0.052
$\frac{16}{100}$	24.8287	−0.066
$\frac{17}{100}$	8.7604	−0.350*
$\frac{18}{100}$	5.2998	−0.164*
$\frac{19}{100}$	5.3633	+0.038
$\frac{20}{100}$	4.6181	+0.058
$\frac{21}{100}$	2.6979	−0.140*
$\frac{22}{100}$	2.4047	−0.109*
$\frac{23}{100}$	3.1497	+0.040
$\frac{24}{100}$	3.3982	+0.034
$\frac{25}{100}$	3.1447	−0.003
$\frac{26}{100}$	2.9615	+0.021

$2T=781\Delta t$ $E\beta(f_j)=-0.008$

$L=50\Delta t$ $D\beta(f_j)=0.052$

Table 6

$f_j\Delta t$	$\hat{p}_{w1}(f_j)$	$b(f_j)$
$\frac{3}{50}$	4.8919	−0.200*
$\frac{4}{50}$	12.0470	−0.299*
$\frac{5}{50}$	33.9834	−0.022
$\frac{6}{50}$	59.4136	+0.125*
$\frac{7}{50}$	53.7521	+0.078*
$\frac{8}{50}$	31.0314	−0.073
$\frac{9}{50}$	17.6957	−0.096*
$\frac{10}{50}$	11.3088	−0.036
$\frac{11}{50}$	6.5494	−0.077*
$\frac{12}{50}$	3.9001	−0.097*
$\frac{13}{50}$	2.7770	−0.082*

$2T=781\Delta t$ $\quad E\beta(f_j)=-0.004$

$L=25\Delta t$ $\quad D\beta(f_j)=0.038$

Table 7

$f_j\Delta t$	$\hat{p}_{w1}(f_j)$	$b(f_j)$
$\frac{6}{100}$	5.0153	+0.030
$\frac{7}{100}$	7.4379	+0.007
$\frac{8}{100}$	9.5385	−0.097
$\frac{9}{100}$	15.5770	−0.084
$\frac{10}{100}$	26.8841	−0.113*
$\frac{11}{100}$	50.7901	−0.002
$\frac{12}{100}$	75.3418	+0.091
$\frac{13}{100}$	71.3490	+0.053
$\frac{14}{100}$	51.7640	−0.038
$\frac{15}{100}$	40.5896	−0.002
$\frac{16}{100}$	29.5078	−0.005
$\frac{17}{100}$	19.1196	−0.091
$\frac{18}{100}$	15.8271	−0.048
$\frac{19}{100}$	15.7318	+0.062
$\frac{20}{100}$	11.5293	+0.003
$\frac{21}{100}$	7.1288	−0.125*
$\frac{22}{100}$	6.5395	+0.044
$\frac{23}{100}$	4.5807	−0.038
$\frac{24}{100}$	3.4551	−0.077
$\frac{25}{100}$	3.3689	+0.056
$\frac{26}{100}$	2.5257	+0.044

$2T=781\Delta t$ $\quad E\beta(f_j)=-0.008$

$L=50\Delta t$ $\quad D\beta(f_j)=0.052$

estimates of power spectral densities were obtained by first taking the (non-circular) mean lagged product, truncating it at the lag L, taking Fourier transform and then averaging by the weight $\{a_n\}$ [1, 3]. It seems that the value $25\times$(unit of time) of L in the examples of tables 4 and 6 is introducing significant biases at the frequencies at and

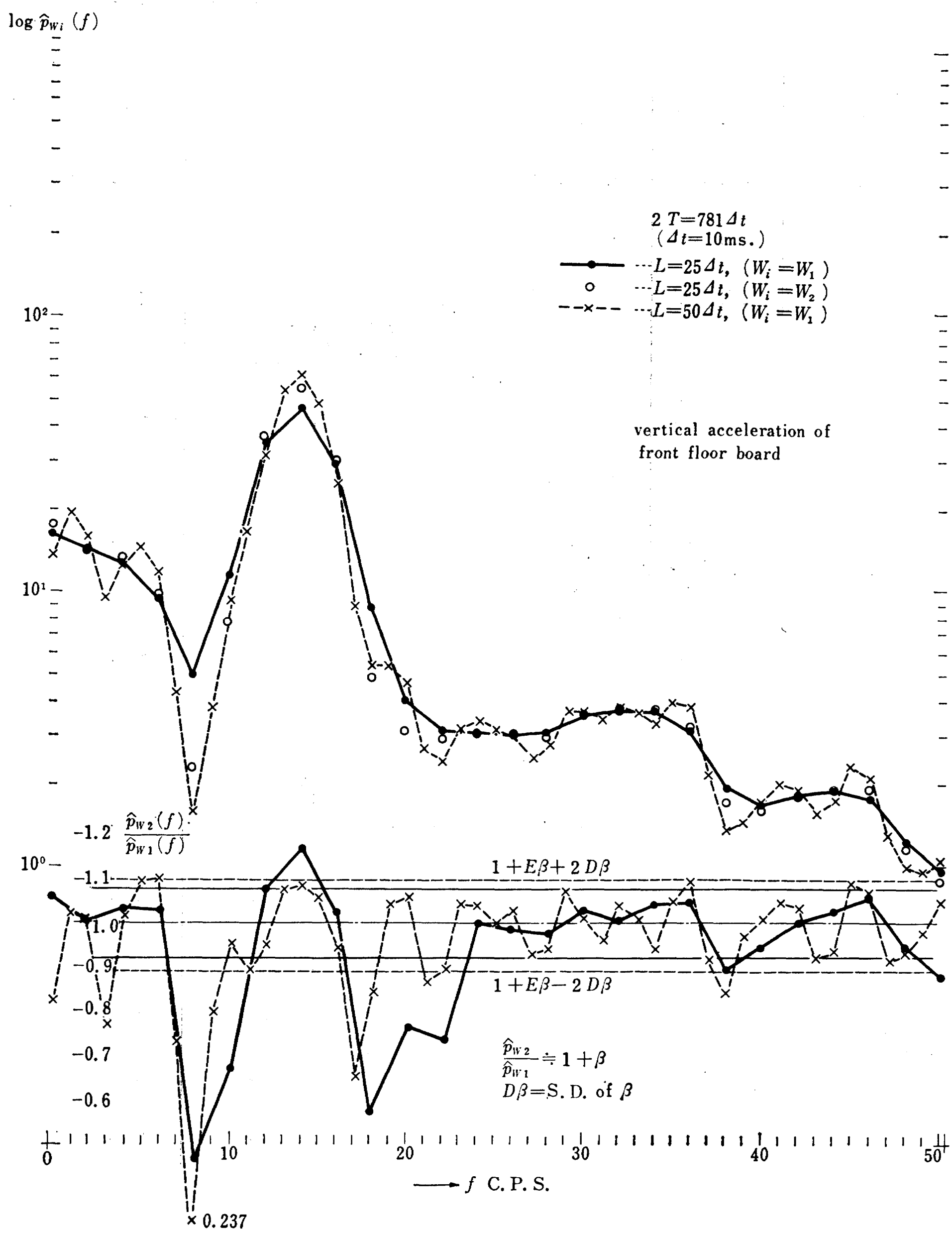

Fig. 1

around the peaks of the power spectra. Taking into account the possibility of highly negative values of $b(f_j)$ under white noise assumption indicated by our sampling experiment, we can see that the value of L for tables 5 and 7, which is twice as large as that of tables 4 and 6, seems to be fairly satisfactory. In fig. 1 are shown the results corresponding to tables 4 and 5. These examples suggest that 2σ criterion, described in the note of table 2, and the knowledge of possible occurrences of highly negative values of $b(f_j)$ provide us with fairly reliable and useful information for the decision of the bandwidth of spectrum window or the length L for each frequency. Of course, such a decision is about hitting a balance between the bandwidth, which is directly concerned with the bias when the shape of the window is fixed, and the variance of the estimate and it would become unnecessary if either one of these were given beforehand.

6. A practical procedure of spectrum computation

It is needless to say that to draw a conclusion from the result of a spectrum computation, we have to use every kind of information about the physical properties of the object under observation and there will never exist a perfectly formal procedure which will tell us about the adequacy of the selection of our window. But our present numerical examples strongly suggest that our $b(f_j)$ computed by using $D_1(t)$ and $D_2(t)$ of section 3 gives a fairly reliable information about the adequacy of the selection of L at various frequencies. We shall here summarize the results in a form which will be directly applicable to practical computations of power spectra.

0. We assume the case where a set of original data is given and the available length of date is fixed.
1. Compute $\hat{p}_{w_1}(f_j)$ and $\hat{p}_{w_2}(f_j)$ for necessary values of f_j.
2. Compute $b(f_j)$ by the formula

$$b(f_j)=\frac{\hat{p}_{w_2}(f_j)}{\hat{p}_{w_1}(f_j)}-1 .$$

3. Compute $E\beta(f_j)$ and $D\beta(f_j)$ which are independent of f_j.
4. Pay attention to the frequencies where

$$|b(f_j)-E\beta(f_j)|>2D\beta(f_j)$$

holds. Especially watch for the specific patterns to be observed at the peaks (and probably at the sharp troughs) of $\hat{p}_{w_1}(f_j)$.

5. When there is a specific pattern in the frequency range of interest and refinement of the estimate is desired, halve the bandwidth

of the spectrum window (or double the length L) for that range.

6. Adopt as our final estimate of power spectrum those $\hat{p}_{w_1}(f_j)$'s which do not suggest the need for any further reduction of the bandwidth of spectrum window and with the largest possible bandwidth or with the smallest possible L.

7. When the expected statistical variability of this final estimate turns out to be too large for the purpose of application, try to get a longer record or to get a set of records of repeated independent observations.

The present problem may be considered approximately to be the problem of detection of a sharp spike like signal in the presence of a back ground noise which is composed of a very low frequency component (smoothed spectrum) and a white noise (sampling fluctuations of periodogram, though the variance changes proportionally to $p(f_j)$). The specific pattern of $b(f_j)$'s at the peak of power spectra is a (smoothed) impulse response of the filter determined by $\{\Delta W_k\}$, which does not respond to the very low frequency variation of $p(f_j)$ in f_j but is acting as a smoothing operator for the (locally) white noise generated by the sampling fluctuation of the original data. It would be an interesting problem to review the design of $\{W_k\}$ and $\{\Delta W_k\}$ from this standpoint also taking into account the use of the method of fast Fourier transform.

Acknowledgements

The present work has been done while the author was visiting Princeton University in the year 1966–1967. The author expresses his hearty thanks to Professor John W. Tukey and other members of the working group of spectrum package for various suggestions and discussions related with the content of the present paper.

Thanks are due to Mr. M. Isida and other members of computing room of the Institute of Statistical Mathematics for providing the data of a random number generator and also to Miss E. Arahata for the necessary computations.

THE INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] H. Akaike, "On the design of lag window for the estimation of spectra," *Ann. Inst. Statist. Math.*, 14 (1962), 1-21.

[2] H. Akaike, "On the statistical estimation of the frequency response function of a system having multiple input," *Ann. Inst. Statist. Math.*, 17 (1965), 185-210.

[3] R. B. Blackman and J. W. Tukey, "The measurement of power spectra from the

point of view of communications engineering," *Bell System Technical Journal*, 37 (1958), 185-282, 458-569. (also published separately by Dover, 1958)

[4] H. E. Daniells, "The estimation of spectral densities," *J. R. Statist. Soc.*, B, 24 (1962), 185-198.